An Application of Genetic Programming for Power System Planning and Operation

R.Behera[1], B.B.Pati[2], B.P.Panigrahi[2], S. Misra
[1, 3] Department of Electrical Engineering
I.G.I.T.Sarang, Orissa; India
b_rabindra@yahoo.co.in
[2, 4] Department of Electrical Engineering
VSSUT Burla, Orissa, India

Abstract: **This work incorporates the identification of model in functional form using curve fitting and genetic programming technique which can forecast present and future load requirement. Approximating an unknown function with sample data is an important practical problem. In order to forecast an unknown function using a finite set of sample data, a function is constructed to fit sample data points. This process is called curve fitting. There are several methods of curve fitting. Interpolation is a special case of curve fitting where an exact fit of the existing data points is expected. Once a model is generated, acceptability of the model must be tested. There are several measures to test the goodness of a model. Sum of absolute difference, mean absolute error, mean absolute percentage error, sum of squares due to error (SSE), mean squared error and root mean squared errors can be used to evaluate models. Minimizing the squares of vertical distance of the points in a curve (SSE) is one of the most widely used method .Two of the methods has been presented namely Curve fitting technique & Genetic Programming and they have been compared based on (SSE)sum of squares due to error.**

Key words: **Power System Planning, Load Forecasting, Curve Fitting, Genetic Algorithm, Mutation, Fitness Function.**

I. INTRODUCTION

One of the primary power system planning tasks of an electric utility is to accurately predict load requirements at all times. Results obtained from load forecasting process are used in different areas such as planning and operation. Planning of future investment for the constructions depends on the accuracy of the long term load forecasting considerably therefore, several estimation methods have been applied for short, mid and long term load forecasting. Conventional load forecasting techniques are based on statistical methods. Stochastic time series, non-parametric regression models were used in load forecasting. Also soft computing techniques were used as load estimator, such as recurrent neural net work of ANN model [8]. The estimation of load in advance is commonly known as Load Forecasting. Power system expansion planning starts with a forecast of anticipated future load requirement. The estimation of both demand & energy requirement is crucial to an effective system planning. Demand predictions are used for determining the generation, capacity transmission and distribution system additions [3]. Load forecasting is also used to establish policies for constructions, capital energy forecast which are needed to determine future fuel requirements. Thus a good forecast reflecting the present and future trend is key to all planning. D.K.Chaturvedi & R.K.Mishra (1995) [3] presented the Genetic Algorithm approach for long term load forecasting. For load forecasting the results obtained through genetic algorithm is compared With the result given by APS(Annual Power Survey) carried out by CEA(Central Electricity Authority).Genetic Algorithms claims to provide near optimal solution or optimal solution for computationally intensive problems.

Dr. Hanan Ahmad Kamal (2002) [4] focused on technique of solving curve fitting problems using genetic programming Curve Fitting problems used to be solved by assuming the equation shape or degree then searching for the parameter values as done in regression techniques. This paper demonstrates that Curve Fitting problems can be solved using GP without need to assume the equation shape. Object oriented technique has been used to design and implement a general purpose GP engine.

M. A. Farahat and M. Talaat (2010) [6] presented a New Approach for Short-Term Load Forecasting Using Curve Fitting Prediction Optimized by Genetic Algorithms Curve fitting prediction and time series models are used for hourly loads forecasting of the week days. It is shown that the proposed approach provide very accurate hourly load forecast. Also it is shown that the proposed method can provide more accurate results. The mean percent relative error of the model is less than 1 %. actual data. The ANN model is then used to forecast the annual peak demand of a Middle Eastern utility up to the year 2006.

Khaled M. EL-Naggar (2005) [2] which presents a paper which describe the comparison of three estimation techniques used for peak load forecasting in power systems. The three optimum estimation techniques are, genetic algorithms (GA), least error squares (LS) and, least absolute value filtering (LAVF). The problem is formulated as an estimation problem. Different forecasting models are considered.

Azadeh, S.F. Ghaderi and S. Tarverdian (2006) [7] presents a genetic algorithm (GA) with variable parameters to forecast electricity demand using stochastic procedures. The GA applied in this study has been tuned for all the GA parameters and the best coefficients with minimum error are finally found, while all the GA parameter values are tested together. The estimation errors of genetic algorithm model are less

DOI: 01.IJCSI.03.02.59

than that of estimated by regression method. Finally, analysis of variance (ANOVA) was applied to compare genetic algorithm, regression and actual data Zargham Hayadri (2007) [1] presented a Time-Series Load Modeling and Load Forecasting Using Neuro-Fuzzy Techniques. In this method, energy data of several past years is used to train an Adaptive Network based on Fuzzy Inference System (ANFIS). ANN structure of ANFIS can capture the power consumption patterns, while the fuzzy logic structure of ANFIS performs signal trend identification

John R. Koza et al.,(1994) [17] presented the survey of genetic algorithm and genetic programming where both method has been compared and their represented scheme of solutions has been deeply focused.

Zhu Huan-rong et al (2010) [18] uses the genetic programming(GP) method to establish the mathematical model of load forecasting to meet certain precision required under the conditions of a particular time in the future developing trend of the load to make estimates and assumptions of science. considering the meteorological factors on the impact of electricity load.

Edgar Manuel Carreno (2011) [21] formulated a paper which forecast a spatial electric load using cellular automation approach The most important features of this method are good performance, few data and the simplicity of the algorithm, allowing for future scalability. The approach is tested in a real system from a mid-size city showing good performance. Results are presented in future preference maps**.**

II. LONG TERM LOAD FORECASTING

This is done for 1-5 years in advance in order to prepare maintenance schedule of generating units, planning future generation capacity, entering into an agreement for energy interchange with neighboring utilities. There are two approaches namely,

A. Peak Load approach

In this simple approach is to extrapolate the trend curve, which is obtained by plotting the past values of annual peak against year of operation. The following analytical function can be used to determine the trend curve [13].

(i) Straight Line $y = a + bx$

(ii) Parabola $y = a + b * x + cx^2$

(iii) Polynomial curve $y = a + bx + cx^2 + dx^3$

y represents peak load and x represents time in years. The most common method of finding coefficient a, b, c, d is the least square curve fitting technique.

B. Energy approach

Another method is to forecast annual sales to different class of customers like residential, commercial industrial etc which can be converted to annual peak demand using annual load factor

III. GENETIC PROGRAMMING

Genetic programming is an extension of the genetic algorithm in which the structures in the population are not fixed-length character strings that encode candidate solutions to a problem, but programs that, when executed, are the candidate solutions to the problem. Genetic programming is a domain-independent method that genetically breeds a population of computer programs to solve a problem. Moreover, genetic programming transforms a population of computer programs into a new generation of programs by applying analogs of naturally occurring genetic operations iteratively .This process is illustrated in Fig 3.1.

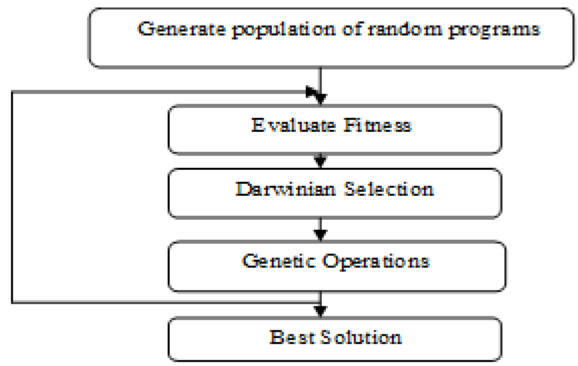

Fig 3.1 Main Loop of genetic programming

A. Genetic Representation

The programs are represented in a tree form in GP, which is the most common form, and the tree is called program tree (or parse tree or syntax tree). Some alternative program representations include finite automata (evolutionary programming) and grammars (grammatical evolution). For example, the simple expression min(x/5y, x+y) is represented as shown in Figure 3.2. The tree includes nodes (which are also called points) and links. The nodes indicate the instructions to execute. The links indicate the arguments for each instruction. In the following, the internal nodes in a tree will be called functions, while the tree's leaves will be called terminals. The trees and their expressions in genetic programming can be represented using prefix notation (e.g., as Lisp S-expressions). A basic idea of lisp programs is required to understand the representations and programming of genetic programming. In prefix notation, functions always precede their arguments. In this notation, it is easy to see the correspondence between expressions and their syntax trees. Simple recursive procedures can convert prefix-notation expressions into infix-notation expressions and vice versa.

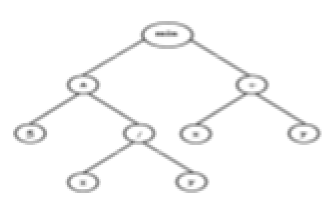

Fig 3.2 Basic Tree-Like Program Representation Used in Genetic Programming

DOI: 01.IJCSI.03.02. 59

The choice of whether to use such a linear representation or an explicit tree representation is typically guided by questions of convenience, efficiency, the genetic operations being used (some may be more easily or more efficiently implemented in one representation), and other data one may wish to collect during runs. These tree representations are the most common in GP, e.g., numerous high-quality, freely available GP implementations use them.

B. Genetic Programming Methodology

Genetic programming starts with a primordial ooze of thousands of randomly created computer programs. This population of programs is progressively evolved over a series of generations. The evolutionary search uses the Darwinian principle of natural selection (survival of the fittest) and analogs of various naturally occurring operations, including crossover (sexual recombination), mutation, gene duplication, gene deletion. In addition, genetic programming can automatically create, in a single run, a general (parameterized) solution to a problem in the form of a graphical structure whose nodes or edges represent components and where the parameter values of the components are specified by mathematical expressions containing free variables. That is, genetic programming can automatically create a general solution to a problem in the form of a parameterized topology.

Preparatory Steps Of Genetic Programming:

Genetic programming starts from a high-level statement of the requirements of a problem and attempts to produce a computer program that solves the problem. The human user communicates the high-level statement of the problem to the genetic programming system by performing certain well-defined preparatory steps. The five major preparatory steps for the basic version of genetic programming require the human user to specify.

i. The set of terminals (e.g., the independent variables of the problem, zero-argument functions, and random constants) for each branch of the to-be-evolved program,

ii. The set of primitive functions for each branch of the to-be-evolved program,

iii. The fitness measure (for explicitly or implicitly measuring the fitness of individuals in the population),

iv. Certain parameters for controlling the run

v. The termination criterion and method for designating the result of the run.

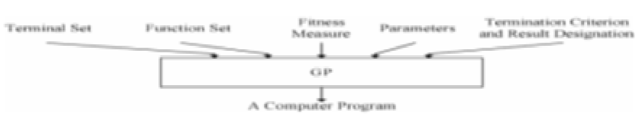

Fig 3.3 Preparatory steps of Genetic Programming

The figure below shows the five major preparatory steps for the basic version of genetic programming. The preparatory steps (shown at the top of the figure) are the human supplied input to the genetic programming system. The computer program (shown at the bottom) is the output of the genetic programming system. The first two preparatory steps specify the ingredients that are available to create the computer programs. A run of genetic programming is a competitive search among a diverse population of programs composed of the available functions and terminals.

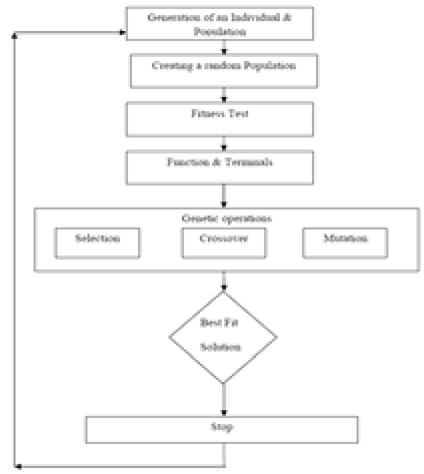

Fig 3.4 Functionality of Genetic Programming

C. Benefits Of Genetic Programming

A few advantages of genetic programming are:

(i) Without any analytical knowledge accurate results are obtained.

(ii) If fuzzy sets are encoded in the genotype, new and more suited fuzzy sets are generated to describe precise and individual membership functions. This can be done by means of the intersection and/or union of the existing fuzzy sets.

(iii) Every component of the resulting GP rule-base is relevant in some way for the solution of the problem. Thus null operations that will expend computational resources at runtime are not encoded.

(iv) This approach does scale with the problem size. Some other approaches to the cart-centering problem use a GA that encodes NxN matrices of parameters. These solutions work badly as the problem grows in size (i.e., as N, increases).

(v) With GP no restrictions are imposed on how the structure of solutions should be. Also the complexity or the number of rules of the computed solution is not bounded

D. Applications Of GP

There are numerous applications of genetic programming. Some of them are:

i. Black Art Problems

ii. Programming The Unprogrammable (PTU

iii. Commercially Useful New Inventions (CUNI)

iv. Optimal Control

IV. RESULTS & DISCUSSION

Accurate load forecasting holds a great saving potential for electric utility corporations since it determines its main source of income, particularly in the case of distributors. Precise load forecasting helps the electric utility to make unit commitment decisions, reduce spinning reserve capacity and schedule device maintenance plan properly. It is therefore

DOI: 01.IJCSI.03.02.59

necessary that the electricity generating organizations should have prior knowledge of future demand with great accuracy. Some data mining algorithms play the greater role to predict the load forecasting. This research work examines and analyzes the use of Curve Fitting Techniques and Genetic Programming (GPLAB) as forecasting tools for predicting the energy demand for three years ahead and comparing the results. Various case studies has been taken from specific areas and energy consumption forecasting has been presented using tools mentioned above.

A. Case Study

Energy consumption for has been taken from Turkey based power utilities started from year 1994 to 2005. Based upon this future energy consumption has been forecasted up to year 2012.In this study, power consumption data is processed with both conventional regression analysis and genetic programming techniques.. Curve fitting tool of MATLAB (cftool) is used for conventional regression and GPLAB Toolbox for MATLAB is used for applying genetic programming. Curve fitting tool of MATLAB can be used to fit data using polynomial, exponential, rational, Gaussian and other equations. It also provides statistics to evaluate the goodness of a fit produced. GPLAB is a free, highly configurable and extendable genetic programming toolbox supporting up-to-date features of the recent genetic programming research. Curve fitting tool is used for comparison with the genetic programming application. Among the different types of the fit, a 4th degree polynomial and a power equation the following form has produced the best results. Coefficients are calculated with 95% confidence bounds. Using Curve Fitting and GP techniques the model found for long term demand forecasting is as follows:

TABLE IV.1 ENERGY CONSUMPTION DATA OF TURKEY BASED POWER UTILITY

Years	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
Mwh	0.583	0.694	0.851	1.071	1.212	1.275	1.395	1.440	1.543	1.569	1.585	1.671

The equation found for 4th degree polynomial model is

$$f(x) = p_1 * x^4 + p_2 * x^3 + p_3 * x^2 + p_4 * x + p_5 \tag{4.1}$$

TABLE IV.2 CALCULATED COEFFICIENTS OF A 4TH DEGREE POLYNOMIAL MODEL

p_1	p_2	p_3	p_4	p_5
2.049*10^5	-5.175*10^6	3.498*10^7	7.127*10^7	4.699*10^8

Table IV.3 Measures of Goodness of fit for polynomial model

SSE	R^2	Adjusted R^2	RMSE
0.00608797	0.9959	0.9935	0.00086971

Here independent variable is taken as year and demand as the dependent variable. Functional form is found which best describes the data while minimizing the error. In conventional regression a model is selected in form of polynomial equation .The above table shows the various measures of checking the model for its accuracy. Here power model is selected, its coefficient is calculated .The year is normalized and is taken from numerical value one and so on. Though power model is nonlinear in nature, it can be converted to linear by taking logarithmic both sides .Though this model suffers from high error rate .This model is selected only to make suitable comparison. While comparing table 5.3 and table 5.7 it is found that SSE of polynomial model is less as compared to power model. Moreover RMSE of polynomial model is better than the power model.

By incorporating the values of independent variable and the calculated coefficient in equations the values of dependent values can be found. Hence the forecasted values of demand can be calculated both for the current and future years respectively.

*Power Model :*The equation found for power model is

$$f(x) = a * x^b + c \tag{4.2}$$

TABLE IV.4 CALCULATED COEFFICIENTS OF POWER MODEL

a	b	c
8.724*10^8	0.3411	-3.424*10^8

TABLE IV.5 MEASURES OF GOODNESS OF FIT FOR POWER MODEL

SSE	R^2	Adjusted R^2	RMSE
0.02194693	0.9852	0.9819	0.003135

*GP Model :*Using Symbolic regression both parameters & Symbolic model is found for long term energy consumption forecasting. For this GPLAB Program has been run for 800 generations and population size of 100. It has fitness value 599396247.98.The function found for symbolic regression at Generation 752 .

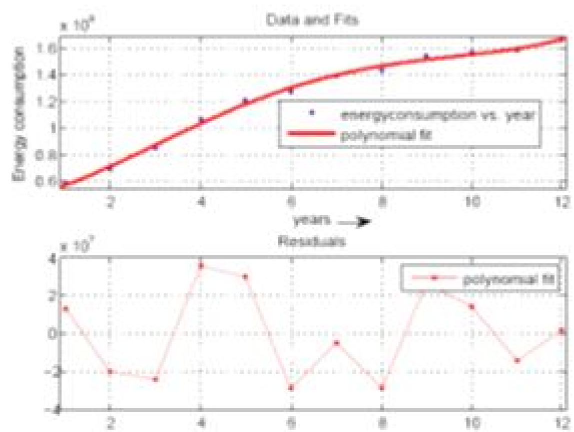

Fig 4.1 Output of polynomial model & it's Residual

DOI: 01.IJCSI.03.02.59

TABLE IV.6 FORECASTING DEMAND FOR TURKEY BASED POWER UTILITY (A COMPARATIVE STUDY)

Years	Output of the Polynomial model(Mwh/10⁹)	Output of the Power model(Mwh/10⁹)	Output of the GP model(Mwh/10⁹)	Error in Percentage		
				Polynomial model	Power model	GP model
1994	0.571	0.529	0.582	0.0216	0.0923	0.0027
1995	0.714	0.762	0.730	0.0288	0.0985	0.00494
1996	0.875	0.926	0.849	0.0282	0.0883	0.0028
1997	1.036	1.057	1.074	0.0332	0.0132	0.0026
1998	1.182	1.168	1.209	0.0247	0.0362	0.0019
1999	1.304	1.265	1.312	0.0226	0.0085	0.0280
2000	1.400	1.351	1.395	0.0032	0.0313	0.0001
2001	1.464	1.430	1.465	0.0200	0.0064	0.0174
2002	1.517	1.503	1.540	0.0171	0.0260	0.0018
2003	1.555	1.571	1.540	0.0090	0.0009	0.0018
2004	1.599	1.634	1.604	0.0090	0.0308	0.0118
2005	1.672	1.693	1.675	0.0010	0.0131	0.0022

TABLE IV.7 FORECASTING OF ENERGY CONSUMPTION FOR NEXT SEVEN YEARS USING (A COMPARATIVE STUDY)

Year	Output of Polynomial Model(Mwh/10⁹)	Output of power model (Mwh/10⁹)	Output of GP Model
2006	1.792	1.750	1.767
2007	1.997	1.803	1.888
2008	2.320	1.854	2.036
2009	2.800	1.903	2.205
2010	3.484	1.950	2.384
2011	4.420	1.995	2.565
2012	5.666	2.039	2.736

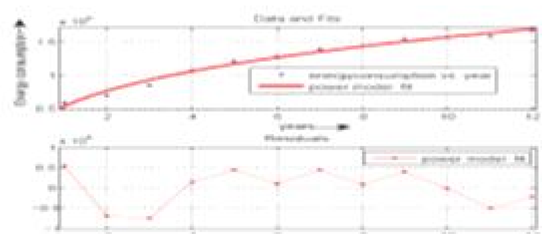

Fig 4.2 Output of power model & its residuals

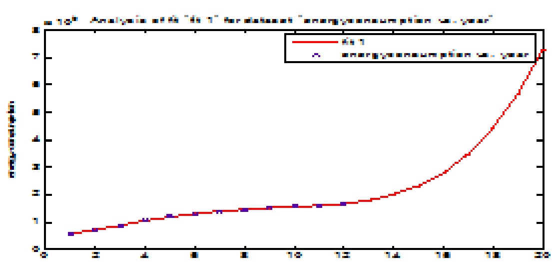

Fig 4.3 Forecasting energy consumption for next year's using polynomial model

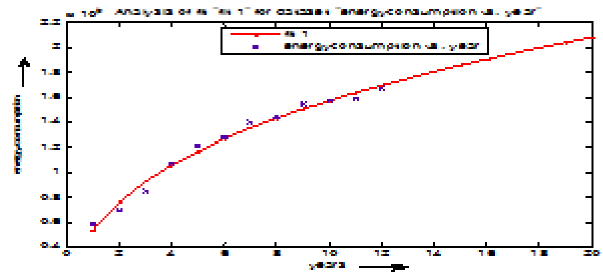

Fig 4.4 Forecasting energy consumption for next year's using power model

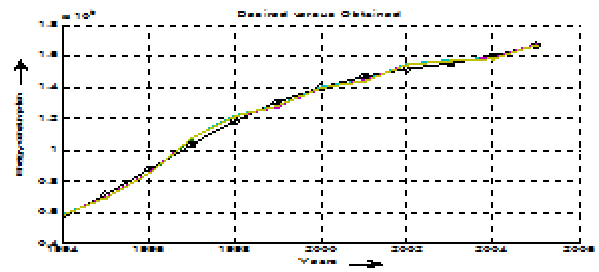

Fig 4.5 Forecasting energy consumption using input data for GP model

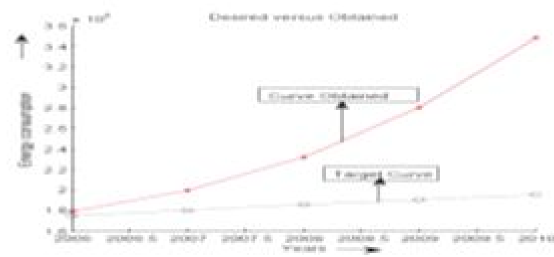

Fig 4.6 Graph for forecasted energy consumption of next seven years using GP model

DISCUSSION

Annual demand data has been taken from various utility as case studies. Both conventional and symbolic regression has been successfully implemented using input data. Considering the demand data of Turkey based power utility, SSE in case of GP model is 0.00038000 which is less than Polynomial model which is 0.0060879 or power model which is 0.02194693.It means if SSE is low the sum of vertical distances between the desire curve and the obtained curve is small, which guarantees the best model. Here GP is run population size of 100 and 800 generations. The best fitness value is found to be 338.50 at generation 792.

CONCLUSION

While forecasting the future energy consumption using the data of turkey base power utilities by the method of curve fitting and genetic programming technique .We have used polynomial model, power model and GP model to forecast the load requirements by providing time as an independent variable, it is found that output produces by power and polynomial model somewhat deviates from the actual data while output produces by GP model closely resembles the actual value and error is less in case of Genetic Programming.

DOI: 01.IJCSI.03.02.59

Moreover SSE nearly approaches zero in case of GP model as compared to curve fitting technique which ensures that model found using GP closely fits the actual data. Hence forecasting of annual energy consumption must be done based upon Genetic Programming which uses symbolic regression technique. This is the advantages of symbolic regression over conventional regression technique .In symbolic regression both model and its coefficients can be found which minimizes the chance of selecting a inbuilt function which may not be a better model as in case of curve fitting technique. To test the model RMSE has been calculated and compared .It is been found to be nearly one in case of GP model while its value deviates from one in case of power model or polynomial model

FUTURE SCOPE OF THE WORK

In the present work Genetic programming is used to forecast the future load requirements which incorporate time as an independent variable and energy consumption as a dependant variable. Practically it may be possible to include weather information, temperature, GDP, Number of consumers as independent variables. We can apply genetic algorithm, ANN, Fuzzy Logic to long term energy forecasting so as to get desired form of accuracy. We can also use GA-ANN and other hybrid optimization technique to forecast the future load requirements.

REFERENCES

[1]. Zargham Haydari and F. Kavehnia (2006) " Time-Series Load Modeling and Load Forecasting Using Neuro-Fuzzy Techniques",9th international conference, electrical power quality and utilization ,page no.9-11.

[2]. Khaled M. EL-Naggar, and Khaled (2005) "Electric Load Forecasting Using Genetic Based Algorithm, Optimal Filter Estimator and Least Error Squares Technique: Comparative Study" World Academy of Science, Engineering and Technology ,volume no. 6.

[3]. D. K. Chaturvedi, R. K. Mishra, and A. Agarwal. (1995) "Load Forecasting Using Genetic Algorithms" *Journal of The Institution of Engineers (India), EL* 76.3, page no.161-165.

[4]. Dr. Hanan Ahmad Kamal (2002) "Solving Curve Fitting problems using Genetic Programming" IEEE MELECON May page no.7-9.

[5]. M. A. Farahat and M. Talaat "A New Approach for Short-Term Load Forecasting Using Curve Fitting Prediction Optimized by Genetic Algorithms (2010) "Proceedings of the 14th International Middle East Power Systems Conference (MEPCON'10), Cairo University, Egypt Volume No.125 December page no.19-21.

[6]. A. Azadeh, S.F. Ghaderi and S. Tarverdian (2006) "Electrical Energy Consumption Estimation by Genetic Algorithm "IEEE ISIE page, no.9-12.

[7]. Edmund, T.H. Heng and Dipti Srinivasan "Short Term Load Forecasting Using Genetic Algorithm and Neural Networks" IEEE Catalogue No: 98EX137 June 1998, page no.21-26.

[8]. Tawfiq Al-Saba and Ibrahim El-Amin (1999) "Artificial neural networks as applied to long-term demand forecasting" Artificial Intelligence in Engineering volume no.13, page no.189–197.

[9]. T. Rashid and T. Kechadi (2005) "A Practical Approach for Electricity Load Forecasting" World Academy of Science, Engineering and Technology, volume no.5.

[10]. A.Azadeh, R. Tavakkoli-Moghaddam,and S.Tarverdian (2010) "Electrical Energy Consumption estimation by Genetic Algorithm and Analysis of Variance.

[11]. Sanjib Misra & S.K.Patra (2008) " Short Term Load Forecasting using Neural Network trained with Genetic Algorithm & Particle Swarm Optimization" First International Conference on Emerging Trends in Engineering and Technology,IEEE computer society.

[12]. Quinlan *et.al* (1990) "genetic programming" paradigm which genetically breeds populations of computer programs to solve problems" Computer Science Department Stanford University.

[13]. K.Timma Reddy 2004 "Forecasting using Neural network and genetic algorithm".

[14]. P.K. Sarangi & R.K.Chauan (2005-2009) "short term load forecasting using neuro genetic hybrid approach".

[15]. R John Koza *et.al* (1990) "genetic programming technique based on Darwinism & natural selection for formulating and solving problems".

[16]. John R. Koza (1994) "Survey of Genetic Algorithm and Genetic progrmming".

[17]. Zhu Huan-rong (2010) "genetic programming(GP) method to establish the mathematical model of load forecasting" International Conference On Computer Design And Appliations

[18]. Limin Huo (2007) "Short-Term Load Forecasting Based on Improved Genetic Programming" IEEE proceeding.

[19]. D. K. Chaturvedi, Sinha Anand Premdayal,and Ashish Chandiok (2010) "Short-Term Load forecasting Using Soft Computing Techniques" Int. J. Communications, Network and System Sciences, volume no.3, page no.273-279.

[20]. Edgar Manuel Carreno 2011 "A Cellular Automaton Approach to Spatial Electric Load Forecasting" IEEE Transactions on power system, vol.26, No. 2.

[21]. Nima Amjady 2011 "Midterm Demand Prediction of Electrical Power Systems Using a New Hybrid Forecast Technique" IEEE Transactions on power system, vol.28, No. 4.

[22]. Nima Amjady 2011 "Midterm Demand Prediction of Electrical Power Systems Using a New Hybrid Forecast Technique" IEEE Transactions on power system, vol.28, No. 4.

[23]. Yang Wang 2011" Secondary Forecasting Based on Deviation Analysis for Short-Term Load Forecasting" IEEE Transactions on power system, vol.30, No. 7.

[24]. Alaa F. Sheta (2001) "Forecasting Using Genetic Programming"IEEE proceeding.

[25]. Yusak Tanoto and Weerakorn Ongsakul (2010) "Long-term Peak Load Forecasting Using LMFeedforward Neural Network for Java-Madura-Bali Interconnection, Indonesia" page no.2-4

DOI: 01.IJCSI.03.02. 59